DOCUMENT RESUME

ED 025 041 EA 001 889

ESEA Title I Projects, Evaluation Report 1967. Volume I. Chapter 1, The Pittsburgh Evaluation Model.
Pittsburgh Public Schools, Pa.
Spons Agency-Office of Education (DHEW), Washington, D.C. Bureau of Elementary and Secondary Education.
Report No-ESEA-Title-1
Pub Date 67
Contract-OEC-1-7-000059-0059
Note-41p.
EDRS Price MF-$0.25 HC-$2.15
Descriptors-*Evaluation Criteria, *Evaluation Methods, Federal Legislation, Guidelines, *Program Evaluation
Identifiers-ESEA Title 1, Pittsburgh

Viewing decision makers within the program as the primary audience for program evaluation information, a model for the evaluation of ESEA Title I projects is developed for program improvement as well as program assessment. The model incorporates four stages of assessment, with five evaluative criteria distinguished for each stage. Both negative and positive criteria define the evaluation cycle. A rationale is given for each phase of the evaluation process, including program definition, stage review, feedback, and problems inventory. (JK)

ESEA

TITLE I PROJECTS

EVALUATION REPORT
1967 - VOLUME I

ED025041

EA 001 889

PITTSBURGH PUBLIC SCHOOLS
SIDNEY P. MARLAND, JR., SUPERINTENDENT

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

ESEA TITLE I PROJECTS

EVALUATION REPORT

1967 - VOLUME I

ED025011

Pittsburgh Public Schools

Sidney P. Marland, Jr., Superintendent

The research reported herein was performed pursuant to contract OEC1-7-000059-0059 with the Office of Education, U.S. Department of Health, Education, and Welfare. Contractors undertaking such projects under Government sponsorship are encouraged to express freely their professional judgment in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent official Office of Education position or policy.

OFFICE OF RESEARCH*

Dr. Malcolm M. Provus, Director
Dr. Russell Scott, Assistant Director
Mrs. Mary Jane Duda, Coordinator of Evaluation
Mr. Gordon A. Welty, Research Associate
Miss Ethel Patton, Research Associate
Dr. Stuart Shaffer, Research Associate
Miss Judith Ann Taylor, Research Assistant
Mr. Seon H. Cho, Senior Systems Analyst
Mr. Forrest Jewell, Computer Specialist

Project Evaluators

Mr. John Calhoun
Mr. Vincent Farley
Mr. Leonard E. Glassner
Dr. Wafik A. Hassouna
Miss Esther Kresh
Mr. Herbert Leifer
Mrs. Stase McPherron
Mr. John V. O'Neill
Mr. Jeffrey PTaschnik
Mrs. Vivien Richman

Graduate Assistants

Mr. Donald Hendershot
Mr. Conway Jeffress
Mrs. Carole Katz
Mr. Varghese Keerikatte
Mr. Frank Pogue
Mrs. Eleanor Smith
Mr. Dan Stern
Mr. Eugene Wallace

Editorial Staff

Mrs. J. H. McBroom, Sr. Editor
Mrs. Yvonne Fogel, Editor
Mr. Carl Bredhal
Miss Linda Craig
Mr. Anthony Crocco
Mrs. Janet McGrail
Miss Alexanne Walgren

Staff Consultants

Mr. Alex Andres
Mrs. Ann Anthony
Dr. Morris Berkowitz
Dr. William Brinckloe
Dr. Garlie Forehand
Dr. Henry Hausdorff
Mr. Alan Klein
Dr. Jerome Laulicht
Dr. Gordon Marker
Dr. Margaret McFarland
Mr. and Mrs. Robert Morey
Dr. Lauren Resnick
Dr. Seth Spaulding
Dr. Edward Suchman
Mrs. Ann Taylor

Clerical Staff

Administrative Assistants:
Mrs. Elizabeth Rodd
Mrs. Jo Anne Tierno

Secretaries and Typists:
Mrs. June Allan
Mrs. Florence T. Browne
Miss Mary Alice Carter
Mrs. Lena M. DiNome
Miss Elaine Florig
Mrs. Esther Gajewski
Miss Jo Ann Glenn
Miss Penny Hase
Mrs. Joan Lasky
Miss Martha J. Menham

*Since this report was the product of task-oriented teams, the titles and classifications shown above refer to the functions served by those named at the time of the evaluation. They do not necessarily reflect the official title of the staff member in the Office of Research.

THE PITTSBURGH EVALUATION MODEL

Background and Context

In Pittsburgh Public Schools a model evaluation plan is being developed under a contract with the United States Office of Education and in conjunction with the evaluation of Title I programs. The proposal for the development of model criteria and procedures in Pittsburgh set forth the following premises:

1. That evaluation in the public school setting is a process for program improvement as well as for program assessment
2. That maximizing the involvement of program personnel (field staff and administrative staff) in the process of evaluation fosters staff commitment to program improvement and promotes desired change in staff behavior
3. That the nondirective role for evaluation personnel is effective in promoting program improvement

In addition, the following conditions or factors in the Pittsburgh setting have contributed to shaping the Pittsburgh Model:

1. Lack of adequate pre-implementation planning and evaluation for Title I programs
2. Anticipation of eventual support for evaluation of other educational programs currently operating within the system

3. Proliferation of Title I programs (28 during the first year of model development), which has made necessary some compromises between the optimum and the practical in developing evaluation procedures

Out of the foregoing considerations and the experience of model building, the following features of the Pittsburgh Model have emerged:

1. The purpose of program evaluation is seen as that of providing information requisite to program development and stabilization and for valid program assessment to those responsible for decisions to change the program.
2. Evaluation and decision-making functions are seen as separate.
3. Educational programs are viewed as discrete subsystems of the total school system.
4. The audience for evaluation information is seen as primarily the decision makers within each program. The exceptions to this rule are at points of inter-program contact (e. g. are separate programs compatible?) and comparison (e. g. which is most efficient?).[1] Decision-making responsibility is

---

[1] In order to satisfy the requirements of the Elementary and Secondary Education Act of 1965 and of the state and federal agencies responsible under the law, product assessment not always consistent with the Pittsburgh model is carried on concurrently with model activities. Product evaluations not consistent with the model are reported annually, whereas evaluation reports under the model are not tied to a fixed time schedule.

presumed to reside in all strata of program staff; i.e., all members of program staff, from teachers on up through the supervisory and administrative ranks, are potential program developers.

5. The focus of evaluation under the model--the evaluation criteria and their sequence--reflects an interpretation of the information input necessary for program improvement and assessment in the public school setting. Program development is seen as a spiraling process, with much recycling of change activity as a program evolves.

6. The procedures of evaluation under the model--the specifications for how evaluation purposes are to be carried out--are engineered to encourage the involvement and commitment of teachers and other categories of field staff to insure acceptance and use of evaluation findings.

The Focus of Evaluation

The first concern of the model building effort in Pittsburgh is that of determining what is to be evaluated. Guidelines have been adopted which relate evaluation to the spiral of program development, specifying both the sequence of evaluation interest over time and the scope of interest in point of time. An explanation of the guidelines and their underlying rationale is presented in the following section.

The Guidelines: Sequence

Guidelines relevant to sequence in evaluation are represented in Figure I. Under these guidelines, evaluation purposes or objectives are seen as related to the decision problems of program developers; these decision problems, in turn, are seen as being determined by the status of the program in the process of program development. For convenience in formulating a generalized plan of sequence, the development process is viewed as being segmented into four levels or stages, and a set of evaluation objectives and criteria is associated with each stage. In Figure I, the four evaluation stages are listed in Column I, and a set of evaluation objectives for each stage in Column II.

Under the guidelines, each evaluation objective is implemented by the evaluation staff through the observation of specified phenomena and the application of specified criteria in making evaluation judgments. To the extent that it is possible at this time to generalize, the guidelines specify the sources of standards for the criteria. In Figure I, the observations and criteria associated with each objective are listed in Columns III and IV respectively. The standards sources for specified criteria are listed in Column V. The development problems for which the evaluation produces information are listed for each stage in Column VI.

Stages of Program Evaluation I	Evaluation Objectives II	Phenomena Observed III	Criteria for Evaluation Judgments IV	Standards Sources V	Development Problems for Decision VI
STAGE I	To assess viability of program design	Public consensus of program staff in response to questions about program dimensions (program design)	Comprehensiveness of program design Face validity of program design	Pittsburgh taxonomy of program dimensions Consultant employed by evaluation staff	Should resources be reallocated within the system? Should the operating program be terminated? Should resources be reallocated within the program (1) to extend the program design or (2) to alter the program design?
	To assess feasibility of program design	Perceptions and opinions of field staff about reciprocal effects of program and program environment	Compatibility of program with program environment	Hierarchy of system objectives implicit in opinions of field staff	
STAGE II	To assess implementation of program and To assess validity of program assumptions	Public consensus of program staff in response to questions about program dimensions (program design) Status of antecedent and process dimensions in operating program	Congruence between operating program and program design in antecedent and process dimensions		Should resources be reallocated to alter the program design and/or to effect operational adjustment (broaden and internalize the public consensus)?
STAGE III	To maintain fidelity of operating program with program design	Public consensus of program staff in response to questions about program dimensions (program design) Status of antecedent and process dimensions in operating program	Congruence between operating program and program design in antecedent and process dimensions		Should resources be reallocated to effect operational adjustment (broaden and internalize the public consensus)?
	To assess effectiveness of operating program and validity of program assumptions	Status of outcome dimensions in operating program	Congruence between operating program and program design in outcome dimensions	Program design relative to criteria for objectives	Should resources be reallocated to alter or restructure the program design?
STAGE IV	To maintain fidelity of operating program with program design	Public consensus of program staff in response to questions about program dimensions (program design) Status of operating program in all dimensions	Congruence between operating program and program design in all dimensions		Should resources be reallocated to effect operational adjustment (broaden and internalize the public consensus)?
	To assess efficiency of operating program and To assess economy of operating program	Time cost of operating program Dollar cost of operating program	Cost relative to effectiveness Cost relative to efficiency	System values System values	Should the program be terminated? Should the program design be altered? restructured?

Figure I
Evaluation Guidelines: Sequence

Stage I. The objectives for evaluation in Stage I are to assess the viability and the feasibility of the program design under the condition that the program is already operational; the paramount question--which definition, or design, of the program to assess--is resolved by the guidelines, which provide for generating a dynamic baseline definition, or working definition, of the program by means of public interviews with representative groups of program staff.[2] This observed public consensus is then judged for viability by applying the two criteria of comprehensiveness and face validity.

The standard used for the criterion of comprehensiveness has been developed in Pittsburgh through systematic analysis of the definitions or designs of all Title I programs. This standard is embodied in a comprehensive list of program elements and is presented in Figure II as the Taxonomy of Program Dimensions. As shown in

---

[2] After a Title I program has been implemented in Pittsburgh, there are usually at least three designs of the program in existence: one is represented by the project proposal, at least one other exists in the thinking of program leaders, and a third is reflected by what actually happens in the operation of the program. Experience suggests that, at any given time, the actual number of designs in existence for a given program is proportional, inversely, to the quality of the intrastaff communication achieved for that program. The rationale for the consensus definition is as follows: it provides an expedient focus for program development activity aimed at modifying and internalizing program goals because (1) it reflects many of the divergent views held by the program staff, while at the same time, (2) it is largely shaped by the thinking of program leaders whose influence in the process of deriving the definition may act as a stimulus for internalization.

GENERAL DESCRIPTION	OUTCOMES	ANTECEDENTS	PROCESS
I. Overall Statement of Objectives and Rationale of the Program II. Description of Scope A. Number of Pupils and/or Number of Classes and Schools Involved B. Grades or Ages of Participants C. General Description of Staff	I. Major Objectives A. Terminal Objectives[1] B. Ultimate Objectives[2] II. Enabling Objectives[3] III. Other Benefits[4] IV. Criteria for Successful Completion of or Removal from the Program	I. Students A. Selection Criteria[5] B. Entering Behaviors[6] II. Staff A. Program Personnel by Specific Positions B. Qualifications for Specific Positions 1. Professional 2. Personal III. Support A. Administrative Support B. Human Resources C. Media D. Facilities IV. Time Constraints	I. Participant Activities A. Enumeration of Activities B. Estimate of Time Spent on Each C. Media Used to Carry out Activities D. Activities Related to Objectives II. Staff Activities A. Functions and Duties for Specific Positions B. Intra-staff Communication and Coordination C. Communication between Program Staff and Others and its purpose

[1] Behaviors exhibited by participants at the end of the program which demonstrate successful completion of the program

[2] Long range goals of the program, objectives to which the program hopefully contributes, but for which it does not have sole responsibility

[3] The skills, attitudes, and information which students must acquire during the program to insure the accomplishment of the major objectives

[4] Benefits expected to accrue to other than program participants as a result of the program

[5] Criteria that are used to determine who shall participate in the program

[6] Characteristics of participants (other than selection characteristics) which are related to performance in the program

Figure II
Taxonomy of Program Dimensions

Figure II, program elements are classified into four broad categories. Three of these--outcomes, antecedents, and process--are regarded as necessary and sufficient for program design. The educational program is conceived as a dynamic input-output system; in keeping with this concept, channels and procedures for communication are seen as essential elements of program design.

The basic questions relative to comprehensiveness of the program definition are (1) "Is there specific program information for each dimension in the taxonomy?" and (2) "Are program dimensions spelled out in acceptable form, i.e., are objectives stated in terms which describe behavior, conditions of behavior, and standards or criteria?" In judging a design for comprehensiveness in the first or second cycles of evaluation,[3] it is common in the Pittsburgh experience to find that many of the dimensions are either not specified, not complete, or not in desired form (e.g. objectives are not stated in behavioral terms). The implication for decision makers in these instances is that curriculum development work or other activity to extend and/or modify the program definition should be initiated.

The second criterion for viability of the program definition is that

---

[3] A description of the evaluation cycle appears in a later section of this paper.

of face validity, or the logic (reasonableness) of the functional relationships implicit in the definition. The standard source suggested by the model for application of this criterion is a qualified consultant employed by the evaluation staff.[4] In judgments relative to face validity reside implications for modification of program design.

To implement the second objective of Stage I evaluation, that is, to assess the feasibility of the program definition by applying the criterion of compatibility, a second set of observations is collected. These observations relate to the reciprocal effects of the program, as currently designed, and the program environment. The question to be answered is "Does this program conflict with other programs or with the school system as a whole in regard to the use of student time, staff time, facilities, or media?"[5]

In the Pittsburgh system, information on which to judge compatibility is not readily available at present. The current practice, therefore, is to assess the perceptions and opinions of field personnel, accepting as a standard the hierarchy of system objectives which is implicit in those opinions. Judgments arrived at in this manner are indeed gross; when unquestionably negative, they lead to questions for

---

[4] A checklist useful for assessing face validity is presented in Figure III. These statements are derived from "Criteria for Stage I Evaluation Judgments" presented in Figure 6 which guides the Pittsburgh evaluation staff.

[5] The basic questions which can help determine compatibility are presented in Figure IV.

	Yes	No
1. The terminal objectives tell what the student does to signify success in the program.	___	___
2. The terminal objectives are clearly related to day-to-day activities of the program.	___	___
3. The enabling objectives are clearly related to the terminal objectives.	___	___
4. The entering behaviors are consistent with the selection criteria.	___	___
5. The entering behaviors are linked to the program's objectives.	___	___
6. The media are clearly related to activities and are sufficient for the essential activities.	___	___
7. The activities are related to the objectives. There is at least one activity for each objective.	___	___
8. Time resources are realistically related to the objectives. There is sufficient time for each specified activity.	___	___
9. Staff qualifications and characteristics are adequate for the functions defined. If not, in-service activities are spelled out to make them adequate.	___	___
10. All functions necessary to serve the objectives are included.	___	___
11. The functions for each group are clearly related to program objectives.	___	___
12. Staff duties are clearly related to staff functions.	___	___
13. Staff duties are clearly defined.	___	___
14. There are duties for each function.	___	___
15. Intra-staff activities are adequate for support of program operation or program objectives.	___	___
16. Communications channels outside the program are related to support needs.	___	___

Figure III
Checklist for Face Validity

1. Is sufficient time available for student participation in the program?
   What activity does the student give up in order to participate in the program?
   Does this reallocation of student time result in sacrifice to other objectives of the school program?
   Does it have an effect on the operation and/or goal attainment of this program?

2. Is sufficient time available for participation by the program staff and cooperating personnel?
   What activities do staff or cooperating non-program personnel sacrifice in order to participate in the program?
   Does this reallocation of their time result in a sacrifice to other objectives of the school program? How does it affect this program?

3. Are facilities and media now available to the program? If not, have plans been made to provide them?
   Is allocation of facilities and/or media to this program resulting in sacrifice of other objectives of the school program?
   Is this program affected by the manner in which facilities and media are allocated?

4. Are the gains for students anticipated by this program equal to, less than, or greater than possible sacrifices in other educational objectives of the school program?

Figure IV
Basic Interview Questions, Program Compatibility

decision making relating to reallocation of resources within the system, program termination, and program modification.

Stage II. The objectives for evaluation in Stage II are to assess both the implementation of the program and the validity of the program assumptions. The initial focus is the current status of the working definition, as reflected by the current public consensus of program personnel. The second set of observations for Stage II evaluation relate to the operational reality of the program in the antecedent and process dimensions.

A single criterion is applied--the congruence or incongruence of current program design with current program operation. This congruence or lack of congruence reflects both on the quality of implementation and/or the validity of program assumptions. If program operation does not accord with intent (as reflected in the public consensus) two, and only two, inferences are possible: (1) functional relationships implicit in the program definition--the program assumptions--are invalid and/or (2) the program has not been communicated and/or internalized among the specified personnel. Implications from these findings relate to action for program redesign and/or for operational adjustment through broadening and internalizing the current consensus.

The guidelines do not yet specify the standard for congruence. The question to be explored is, "How much discrepancy between program design and program operation should be regarded as inevitable?"

Perhaps the standard which evolves will be a ranking of antecedent and process dimensions in terms of the relative importance of discrepancy in each dimension.

Stage III. The observations and judgments which implement Stage II objectives serve another purpose at Stages III and IV. At these higher stages in the sequence, evaluation serves both to monitor the quality of implementation, and as stated in the guidelines, to maintain the fidelity of the operating program with the program design. Thus, at Stage III the working definition is compared with the observed status of antecedent and process dimensions in the operating program. Any resulting implications for change activity relate to operational adjustment through improved communication of the program.

The other purpose of evaluation at Stage III is assessment. Both the effectiveness of the operating program and the validity of program assumptions are assessed in one set of operations--the observation of program outcomes and comparison of these for congruence with outcomes specified in the design. The source of standards for judging congruence is identified by the guidelines as the program's own design (working definition). As seen in the Taxonomy (Figure II) one necessary element of program design is the specification of criteria for outcomes.

Inferences inherent in negative assessment at Stage III relate to validity of the program design. Instability as an explanation of failure

has been ruled out at this stage by both the accomplishment of evaluation objectives at Stage II and the subsequent quality control function of evaluation. Therefore, when it is found that the operating program is not achieving its objectives, the validity of functional relationships implicit in the program definition is called into question. The implication for program leaders is change activity directed toward altering or restructuring the program design.

Stage IV. The quality control function of evaluation continues in this stage and is represented in the guidelines as the first objective for Stage IV evaluation--to maintain the fidelity of the operating program with the program design. As in Stage III, this objective is implemented by first observing both the current working definition of the program and the status of the operating program and then applying the criterion of congruence. Implications from negative evaluation relate to operational adjustment through improved communication of the program.

The unique objectives of evaluation in Stage IV are to assess the efficiency and the economy of the operating program. The observation relevant to assessing the efficiency of the program is its cost in terms of student or participant time (one aspect of the antecedent dimension). The criterion of cost relative to effectiveness is applied, invoking standards derived from and for the entire school system.[6]

[6] These standards do not yet exist in the Pittsburgh system.

Implications from negative assessment in Stage IV relate to decisions for program termination or program modification.

The Guidelines: The Evaluation Cycle

The guidelines for sequence are purposefully inexplicit about the time dimension in evaluation. The range of evaluation interest in relation to periods of time during program evaluation is clarified in the explanation of the evaluation cycle--the mechanism for adapting the evaluation sequence to the rate of change in the program. Under the guidelines, a cycle of evaluation consists of those evaluation activities essential to implementing a set of evaluation objectives and, in addition, a monitoring of program change at the end of the cycle. The evaluation objectives which define an evaluation cycle are not identical, necessarily, with any single group of objectives that defines an evaluation stage (see Figure I).

After the first cycle, each succeeding evaluation cycle is defined in part by the necessity of evaluating in the light of changes in the program. Thus, as program development activity recycles, so also does evaluation activity. As successive evaluation cycles adopt new objectives moving ahead through the sequence of evaluation stages, they also include some of the objectives and associated activities of earlier stages. After Cycle I, an evaluation cycle is likely to be broader in its scope of interest than a single evaluation stage.

Criteria for Definition of the Evaluation Cycle. Experience with the interaction of evaluation activity and program change currently is insufficient to suggest more than general guidelines for cycle definition. These are as follows:

1. Negative Evaluation--When a negative evaluation judgment is reached during one cycle of evaluation, the associated criteria and activities are included again in the next cycle of evaluation. New objectives associated with a higher stage of evaluation are incorporated as long as consistent with the actual progress of program change. There is always a time delay between evaluation judgments directed toward program change and actual program change. Furthermore, pursuant to negative judgment at a given level of program evaluation, program development must recycle at this level and possibly also at earlier levels before "catching up." The decision to define a new evaluation cycle in terms of the next higher level of evaluation is based on expectation that program recycling will be rapid enough for sufficient progress within the time constraints imposed by the structure of a new evaluation cycle. Considerations which affect the expectation of adequate program progress are as follows:

a. Observation of the change decisions and plans of program leaders and judgments of the efficacy of these plans by those members of the evaluation staff most knowledgeable about the program. (Is the

Criteria for Definition of the Evaluation Cycle. Experience with the interaction of evaluation activity and program change currently is insufficient to suggest more than general guidelines for cycle definition. These are as follows:

1. Negative Evaluation--When a negative evaluation judgment is reached during one cycle of evaluation, the associated criteria and activities are included again in the next cycle of evaluation. New objectives associated with a higher stage of evaluation are incorporated as long as consistent with the actual progress of program change. There is always a time delay between evaluation judgments directed toward program change and actual program change. Furthermore, pursuant to negative judgment at a given level of program evaluation, program development must recycle at this level and possibly also at earlier levels before "catching up." The decision to define a new evaluation cycle in terms of the next higher level of evaluation is based on expectation that program recycling will be rapid enough for sufficient progress within the time constraints imposed by the structure of a new evaluation cycle. Considerations which affect the expectation of adequate program progress are as follows:

a. Observation of the change decisions and plans of program leaders and judgments of the efficacy of these plans by those members of the evaluation staff most knowledgeable about the program. (Is the

impetus and know-how for change activity adequate? Is the available time adequate?)

b. Nature of the evaluation observations needed at the higher level of evaluation and time constraints affecting the collection of such observations. (Can program design changes be completed and monitored in time to determine the source of observations and to make collections within the duration of the ensuing cycle?)

2. Positive Evaluation--When positive, or predominantly positive, judgments are reached during a cycle of evaluation, the ensuing cycle is defined primarily by objectives and activities associated with the next higher stage of evaluation.

The Procedures of Evaluation

Guidelines for sequence and cycling define the focus of evaluation. Equally important in the Pittsburgh Evaluation Model are the procedures for implementing these guidelines. The procedures currently employed for making and reporting the specified observations and judgments are described below. For the most part, these apply for Stage I evaluation (project development activity is presently focused on Stage II procedures), although their potential for use in subsequent stages is apparent. The premises of maximum participation of program personnel and nondirectiveness on the part of evaluation person-

nel have played a prominent role in defining the evaluation procedures herein described.

The Program Definition Meeting: Rationale

At the start of evaluation, Title I programs are ill-defined and poorly communicated. Lack of agreement on program objectives is reflected in the distinctiveness of the goal descriptions from different sources. Typically, at least three distinct sets of goals are presented: one set by the project proposal; a second set by program leaders; the third and additional sets of goals, by implication, in the field operation of the program. Deficiencies of program design are evidenced in lack of specificity regarding the antecedent conditions and procedures for reaching objectives and in the vagueness of the objectives themselves.

Whereas the confusion of goals and deficiencies of design suggest the emphasis for evaluation in the initial stages, they also represent a substantial procedural barrier: whose objectives and whose definition of the program do we evaluate?

The solution is provided by the Program Definition Meeting, a procedure of structured group interview for eliciting and recording the public consensus of the entire program staff (or a representative sample) on questions relative to program dimensions. This procedure encourages maximum participation of program personnel in the pro-

cess of deriving a baseline, or working definition, of the program. The product of this meeting--the working definition--reflects many of the divergent views held by program staff, while at the same time it is largely shaped by the thinking of program leaders whose influence may act as a stimulus for internalization. Thus, it is a convenient focus for program development activity aimed at modifying and internalizing program goals.

The Program Definition Meeting: Description

A. Purpose

1. To generate a definition of the program which
   a. Is as complete as possible in terms of the Pittsburgh Taxonomy of Program Dimensions (See Figure I)
   b. Reflects divergent opinion within the program staff
2. To encourage communication between levels and within levels of program staff and to increase knowledge about the program
3. To encourage acceptance of the value of evaluation
4. To maintain a nondirective role for evaluation staff

B. Participants

1. Program staff
   a. When possible, the entire program staff, including the chief program administrator and his superior in the administrative ranks, is invited to the meetings.

b. When participation of total staff is not possible, a stratified sample of staff is invited. Variables taken into account are those which relate to variability of program goals such as (1) length of service in the program, in the school, and in the profession; (2) the size of the school; (3) socio-economic characteristics of the school; and (4) function in the program.

2. Program staff participants enter the meeting with varying degrees of interest in and knowledge about the program. Many, especially at higher levels, are apprehensive about the effects of evaluation. Many at lower levels are reluctant to express divergent opinion in the presence of program leaders.

3. Evaluation Staff

   a. Discussion leaders: Members of the field research staff with leadership experience and/or training in theory of group process and in techniques of discussion leadership play a major role in the definition meeting. These leaders are briefed one or two days before the meeting by the program evaluator.

   b. Status leaders: The Director of Research and/or the Coordinator of Evaluation are present at the meeting.

c. Program evaluator: The member of the evaluation staff with major responsibility for evaluation of the program is involved before, during, and after the meeting.

C. Other Resources

1. A large meeting room with space and facilities for subdivision of the participants into a series of small (maximum of 10 participants) discussion groups is used. The maximum number of small groups accommodated is six. Thus, the total number of participants per meeting is limited to 60.

2. One half-day is the limit of time available to members of the instructional staff for participation in the Program Definition Meeting. Definition meetings adapted to this constraint are conducted for a three-hour morning or afternoon session.

D. The Process

1. Planning and coordinating the meeting is the responsibility of the program evaluator. Prior to the meeting he completes the following plans and preparations:

a. Arrangement of details of date, time, and location of meeting and selection and inviting of participants. In all of these activities the program evaluator works closely with the program administrator or other program leader.

b. Preparation and distribution of the discussion agenda. Dimensions of the program as specified in the Taxonomy of Program Dimensions are interpreted to program personnel in the form of questions about their program. The program evaluator phrases these questions in terms familiar to program personnel. Copies of the discussion agenda, or interview schedule, are distributed to the participants in advance of the meeting.

c. Assignment of participants and discussion leaders to groups. Specific assignment of participants serves several purposes:

(1) To represent within each group those variables--function, attitudes, status, length of experience, physical location--which relate to variability in program goals

(2) To set the stage for maximum input by each individual in the group through strategic placement of authority figures and prevention of interpersonal status barriers

(3) To limit the size of groups to a maximum of ten members each

d. Briefing of discussion leaders. At least one day before the meeting, the program evaluator conducts a briefing for

discussion leaders during which the following kinds of information are presented:

(1) History of the program's development and implementation

(2) Rationale and general description of the program

(3) History of evaluation of the program

(4) Interpretation of the interview schedule

(5) Composition of individual discussion groups, including identification of variables relevant to the discussion process known by the evaluator to be present within the groups--role or value conflicts, status sensitivities, attitudes of program personnel, etc.

(6) Details of physical arrangements, time allocation, and other administrative matters

2. The meeting is opened by status leaders. Leaders from outside the evaluation staff are encouraged to make opening remarks and to lend their support to the goals of the meeting. The status leader representing the evaluation staff explains the nature and purpose of the meeting, including its place in the evaluation process, and stresses the value and importance of individual contributions in the process of program definition. No more than fifteen minutes is devoted to this part of the program.

3. The participants are assigned to small groups for the interviewing. The discussion leader for each group explains again the purpose of the group session and the details of produre:

   a. He stresses the equality of participants in the process and the value of every contribution.

   b. He describes the method of circular response, the procedure for reaching the goal of equal and/or maximum individual participation.

   c. The leader may use a technique such as introduction of individual group members as a means of establishing a favorable climate for the work of the group.

4. The discussion leader uses the interview schedule and the method of circular response to interview the group. The leader continually interprets, feeds back, and summarizes the responses of participants in order to record group consensus for each question. When consensus is not possible, division of opinion is recorded. (The leader may appoint a recorder to supplement his own note-taking.) The interview continues for two hours or more.

5. The program evaluator monitors the discussion groups and decides when the group interview session can be terminated.

At his signal, participants reconvene as a unified group for concluding statements by the program evaluator or the status leader. The leader describes the use to be made of the group interview data and promises feedback to the participants in the form of a formal definition of the program.

6. The program evaluator conducts a post-meeting session of group leaders for discussion and analysis of the interview processes and the resulting products. The evaluator gains insights which help him in the analysis and synthesis of the group interview data and the preparation of a formal definition of the program.

7. The evaluator prepares a formal definition of the program which is then distributed to all participants or to all members of the program staff.

Stage I Panel Meeting: Rationale

A danger inherent in judgmental evaluation is subjectivity in the choice and application of judgment criteria. Alternatives for either avoiding or mitigating the effects of this danger are (1) to confine evaluation to information gathering without judgment, or (2) to employ only criteria identified by the interested parties. The Pittsburgh Evaluation Model rejects these alternatives as inefficient for the purpose of stimulating program improvement.

To preserve the judgmental function of evaluation while minimizing the inherent danger, the Pittsburgh Model relies on two strategies. First, evaluation criteria are predetermined, universally applicable (that is, applicable to all programs evaluated), and external to the individuals responsible for applying them. Secondly, on the presumption of greater safety in numbers, the model employs the panel meeting as the procedure for making evaluation judgments.

Stage I Panel Meeting: Description

A. Purpose

1. To make judgments of the program definition by applying the Stage I criteria
2. To encourage acceptance and use of findings by program decision makers
3. To maintain a nondirective role for evaluation staff

B. Participants--Members of the Panel

1. The program administrator, who may enter the panel meeting somewhat apprehensive about the role and effects of evaluation
2. The program evaluator, the member of the evaluation staff with major responsibility for evaluation of the program who is involved before, during, and after the meeting
3. A consultant employed by the evaluation staff who is a content

specialist in the area of the program's major emphasis

4. A resource person from the evaluation staff who is technically competent in matters of instructional objectives and program design

5. A status leader from the evaluation staff, usually the Coordinator of Evaluation

C. Other Resources

1. Facilities and Equipment

The desired atmosphere is best achieved in a small, comfortable room containing a round table with capacity for seating five persons. A tape recorder operates throughout the meeting to record the deliberations and findings of the panel for subsequent analysis.

2. Materials

Before the meeting, copies of the Program Definition, product of the Program Definition Meeting, and the Taxonomy of Program Dimensions (See Figure II) are sent to each participant. The program administrator and the consultant are given, in addition, materials which describe the purposes and procedures of the panel meeting (See Figure V, Guidelines for Stage I Judgments). During the meeting, the panel moderator is guided by these procedures and by the Criteria for Stage I Evaluation Judgments (See Figure VI).

1. Judgments of the Program Definition will be made by a panel consisting of the program evaluator, the program director, one consultant employed by the Office of Research, and one resource person from the Office of Research, with the Coordinator of Evaluation acting as moderator.

2. The criteria and standards to be applied are as follows:

Criterion	Standard
a. Comprehensiveness of the Program Definition	Taxonomy of Program Dimensions
b. Face validity (internal consistency) of the Program Definition	Consultant employed by evaluation staff
c. Compatibility of defined program with program environment	Values implicit in opinions of field staff

3. The Program Definition and Taxonomy of Program Dimensions are supplied to panel members in advance of the panel meeting. Field observations are reported by the evaluator at the meeting of the panel.

4. In the event that the panel does not agree in making any one judgment, the following rules are to be applied by the moderator:

   a. The evaluator must be responsible for judgments based on criterion "a" above.

   b. The consultant and the resource person from the Office of Research must be responsible for judgments based on criterion "b" above.

   c. The program director and evaluator must be responsible for judgments based on criterion "c" above.

5. A detailed record of the judgments of the panel is to be kept by the evaluator and used as the basis for a Stage I evaluation report to the program staff.

Figure V
Guidelines for Stage I Judgments

Program Definitions are weighed with the three following criteria in mind: (1) comprehensiveness, (2) face validity, and (3) compatibility.

Comprehensiveness

The basic questions here are (1) "Is there specific program information for each dimension in the taxonomy?" (Gaps should be indicated.); and (2) "Are program dimensions spelled out in acceptable form, i.e., are objectives stated in terms which describe behavior, conditions, and standards or criteria?"

Face Validity

Here the internal consistency of the definition is examined through consideration of the following questions:

1. Do the terminal objectives tell what the student does to signify success in the program?
2. Are the terminal objectives clearly related to day-to-day activities of the program?
3. Are the enabling objectives clearly related to the terminal objectives?
4. Are the entering behaviors consistent with the selection criteria?
5. Are the entering behaviors linked to the program's objectives?
6. Are the media clearly related to activities? Are they sufficient for the essential activities?
7. Are the activities related to the objectives? Is there at least one activity for each objective?
8. Are time resources realistically related to the objectives? Is there sufficient time for each specified activity?
9. Are staff qualifications and characteristics adequate for the functions defined? If not, are in-service activities spelled out to make them adequate?
10. Are all functions necessary to serve the objectives included?
11. Are the functions for each group clearly related to program objectives?

Figure VI
Criteria for Stage 1 Evaluation Judgments

12. Are staff duties clearly related to staff functions?

13. Are staff duties clearly defined?

14. Are there duties for each function?

15. Are intra-staff activities adequate for support of program operation or program objectives?

16. Are communications channels outside the program related to support needs?

Compatibility

In measuring the program's compatibility, the question to be answered is: "Is this program consistent with other programs and with the entire school system?" The basic Stage I interview questions can help determine the external consistency of the program.

1. Is sufficient time available for student participation in the program? What activity does the student give up in order to participate in the program? Does this reallocation of student time result in sacrifice of other objectives of the school program? Does it have an effect on the operation and/or goal attainment of this program?

2. Is sufficient time available for participation by the program staff and cooperating personnel? What activities do staff or cooperating nonprogram personnel sacrifice in order to participate in the program? Does this reallocation of their time result in a sacrifice of other objectives of the school program? How does it affect this program?

3. Are facilities and media now available to the program? If not, have plans been made to provide them? Is allocation of facilities and/or media to this program resulting in sacrifice of other objectives of the school program? Is this program affected by the manner in which facilities and media are allocated?

4. Are the gains for students anticipated by this program equal to, less than, or greater than possible sacrifices in other educational objectives of the school program?

5. In general, what is the effect of this program on the environment in which it operates? What is the effect of the environment upon the program?

Figure VI
Criteria for Stage I Evaluation Judgments (contd.)

D. The Process

1. Planning and arranging the meeting is the responsibility of the program evaluator. Prior to the meeting, he completes the following plans and preparations:

a. Arrangement of details of date, time, and location of meeting and selection and inviting of participants. In choosing the consultant for the panel meeting, the evaluator confers with the Coordinator of Evaluation. Three considerations influence the choice of the consultant:

(1) Preparation in the discipline or field of study related to the program's major emphasis

(2) Personal capacity for a productive relationship with program leaders

(3) Opportunity for a continuing relationship with program leaders, as affected by factors such as permanence of professional or vocational ties

b. Preparation of participants before the meeting. At least one day in advance, the evaluator distributes the materials of the meeting and explains its purposes and procedures and the roles of participants to the consultant and to the program administrator. In communicating with the program administrator, he stresses the sharing of responsibility for evaluation judgments.

2. The meeting is opened by the status leader from the evaluation staff, who keynotes the discussion as shared responsiblilty on the part of evaluation staff, program staff, and the consultant for evaluative analysis of the program definition.

3. The status leader acts as moderator for the panel proceedings.

   a. He raises questions relative to the comprehensiveness of the program definition (see Figure VI). In the event of disagreement among the members of the panel on any one question, the moderator invokes the judgment of the program evaluator.

   b. He raises questions relative to the face validity of the program definition (see Figure VI). In the event of disagreement among the members of the panel on any one question, the moderator invokes the combined judgment of the consultant and the resource person from the evaluation staff.

   c. He calls upon the program evaluator to present the results of the Stage I field interviews. He asks for the panel's judgment of program compatibility. In the event of disagreement among the members of the panel, the moderator invokes the combined judgment of the program administrator and the program evaluator.

4. The program evaluator uses the tape recording of panel

proceedings to prepare a comprehensive outline of panel findings, relating them, item by item, to specific parts of the program definition. These findings are the culmination of Stage I evaluation and form the basis for the Stage I Evaluation Report.

5. The evaluator prepares a report of Stage I evaluation which is then distributed by the Office of Research to all members of the program staff.

Feedback: Rationale

The function of evaluation in program development and program assessment under the Pittsburgh Model is to provide both judgmental and objective information about the program for use by program decision makers. It is important that this feedback be given in a manner which encourages both the acceptance of evaluation and the use of evaluation findings for program development. Some of the guidelines which shape the form, focus, timing, and frequency of feedback are as follows:

1. Feedback is always given in a context which provides interpretation of evaluation functions and activities.
2. Feedback is given as promptly after each act of evaluation (each set of observations or judgments) as is consistent with care and accuracy of data handling. It is given in oral form

while written reports are in preparation.

3. The courtesy of a preview of each written report, before it is issued, is extended to the program administrator.
4. For every program, written evaluation reports are issued to all individuals identified as program staff or as resource personnel essential to the program.
5. In written reports, evaluation findings are presented as judgments with implications for action.
6. Interview data is returned to respondents in written form, for preview as well as for verification, before it is used for evaluation purposes.
7. Relevant feedback is given to all adult individuals within the system, whether or not on program staff, who have contributed to evaluation as subjects of observation or participants in evaluation activity.

Feedback: Description

Informal. The evaluator interacts almost continuously with program staff, both administrative and field personnel, the degree of interaction being determined both by the size and scope of the field to be covered and by the number of scheduled activities. Program activities such as in-service training meetings and group planning sessions, as well as the

evaluation activities described in earlier parts of this paper, provide opportunity for informal contacts. Between scheduled activities, the evaluator visits the field operations making personal contacts with individual members of the field staff.

During these visits the evaluator seizes every opportunity for communicating recent evaluation findings and observes and records the reactions of program staff. He is systematic in directing feedback to those individuals who have given time and effort for evaluation. His contacts are most frequent with the program administrator to whom he reports not only the results of evaluation, but also the reactions of field personnel.

Formal. Feedback of this type is provided by the documents described below:

1. Evaluation Reports--At the end of each cycle of evaluation for each program, a written report is issued, with evaluation judgments and any implications for program change activity made clear. The report includes the data on which judgments are based. One section is devoted to an explanation of the nature and purpose of evaluation for the cycle. Since evaluation cycles are not tied to a specific time schedule, evaluation reports may be written as often as several times during a given school year.

2. The Annual Report--The annual report for each program satisfies the requirements of the state and federal agencies responsible

under the Elementary and Secondary Education Act of 1965 for the evaluation of Title I projects. In response to the needs and wishes of these agencies the report includes program assessment, whether or not such assessment is consistent with the evaluation model. (In anticipation of the annual evaluation report, product data is collected concurrently with process data during the school year.) In addition, the report includes a review of all evaluation activity and findings of the previous year, thus providing a view of the individual cycle evaluations as links in the spiral of program development.

3. Monthly Newsletter--As an additional means of insuring continuous feedback to all program staffs, the evaluation staff plans to institute a monthly evaluation newsletter for all personnel involved in Title I programs. The newsletter will describe (1) plans for evaluation as they evolve for each program; (2) current evaluation activities such as program definition or redefinition meetings, panel meetings, instrument development work, data collection, etc; and (3) evaluation findings. One anticipated benefit of the newsletter is the enriching of the concept of evaluation held by field personnel.

The Problems Inventory: Rationale

In Stage II, evaluation looks for lack of congruence between the defined program and the operating program. In thus leading the way to improved communications within the program staff and/or to program

redefinition, evaluation functions for program development. However, a hazard to efficiency in stimulating program change activity is present in the number and range of dimensions for each program and in the current scarcity of instruments for measuring antecedent and process dimensions. For a given cycle of evaluation, thoroughness in seeking out incongruence for every program dimension can be achieved only at the cost of delay in pointing the direction of program improvement. To maximize the benefits of rapid program recycling, evaluation strategy for Stage II relies on selectiveness which is supported by a means for cultivating sensitivity to areas of probable incongruence. The problems inventory is a set of cues which facilitates the assigning of priorities for Stage II observations.

The Problems Inventory: Description

The problems inventory is developed at any convenient gathering of the program staff, or a representative sample of the staff. (Usually time is set aside for developing the problems inventory near the end of a program definition meeting). The inventory is made up of staff responses to a question such as "What single problem in the conduct of your program is most threatening to the objectives of the program?" When conditions permit, small group discussions are used to evolve a ranking for the items on the inventory.

Other Procedures

Project activity currently is aimed toward developing procedures to implement evaluation in Stage II and subsequent stages and, in addition, to facilitate evaluation staff--program staff communication and collaboration across all stages of evaluation. Specifically, procedures are being designed for:

(1) support of program staff activity for refining statements of objectives and the internal structure of program design;

(2) using field personnel for instrument development;

(3) imparting to program personnel a deeper understanding of the evaluation function and of the value to both evaluation staff and program staff of frequent interaction and open channels of communication.